# CSE 150A 250A
# AI: Probabilistic Models

Fall 2025

Instructor: Trevor Bonjour

*Slides adapted from previous versions of the course (Prof. Lawrence, Prof. Alvarado, Prof Berg-Kirkpatrick)*

# Agenda

- Course Logistics
- Syllabus Review
- Course Overview
- Probability Review

# Course Website

https://tbonjour-courses.github.io/cse150a250a-fa25/

We are here to help!

It’s a BIG class! We need to work together!

# Course enrollments are controlled by the department.

- They don’t let me as an instructor directly approve EASY requests.
- Please submit one – department will try to approve as many as they can.
- The course is limited by TA capacity and room size. Can’t add more seats.

I have another class right after this, so I won’t be able to stick around after the lectures :/

# Prerequisites

- Programming
  - Most HW assignments will involve coding in **Python**.
  - Also, basic data analysis and visualization.
  - We can help with algorithmic and conceptual issues.
  - We cannot help with installing, compiling, plotting, etc.

Non-CS backgrounds are welcome.

# Prerequisites

- Elementary probability:
  - Random variables — discrete and continuous
  - Expected values (via sums and integrals)
- Multivariable calculus:
  - Chain rule
  - Gradients and partial derivatives
  - Computing maxima and minima
  - Constrained optimization with Lagrange multipliers

# Prerequisites

- Linear algebra
  - Vectors and matrices
  - Matrix multiplication, inverses, determinants
  - Systems of linear equations
- Mathematical maturity
  - Patience and persistence go a long way
  - Willingness to fill in gaps

# Course Overview

- What we do cover:
  - Inference and learning in Bayesian networks
  - Markov decision processes for reinforcement learning (RL)
- What we don’t cover (not exhaustive):
  - Neural architectures (though we will talk about deep learning)
  - Purely logical reasoning
  - Heuristic search (A*)
  - Theorem proving
  - Genetic algorithms Philosophy of AI

# Course Overview

- What we do cover:
  - Inference and learning in Bayesian networks
  - Markov decision processes for reinforcement learning (RL)

Why these topics?

Skyscraper

Modern AI
(ChatGPT, LLMs)

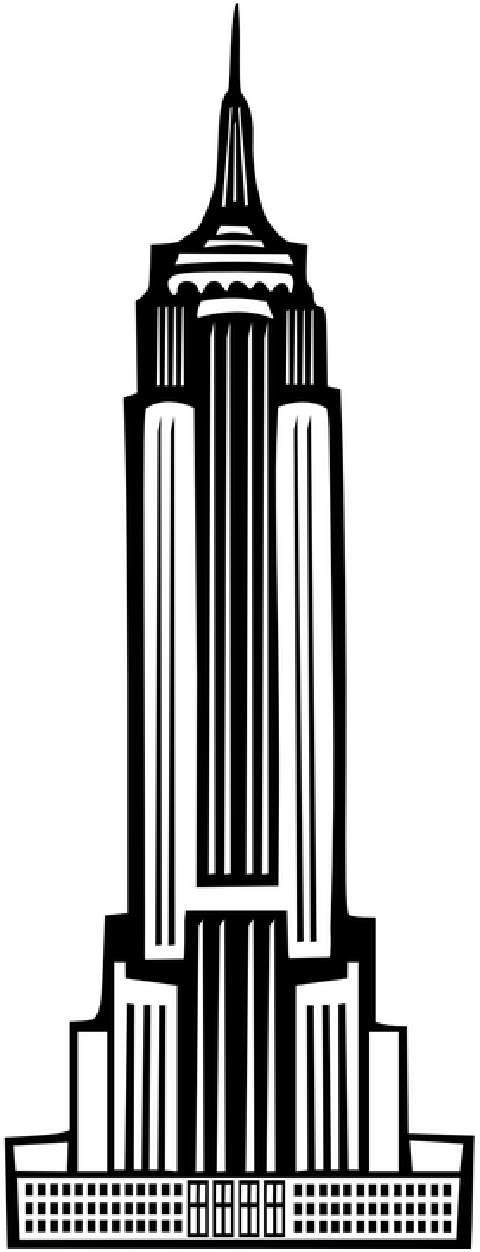
Beams/Columns

Neural Networks
(Deep Learning)

Foundation

**(THIS COURSE)**

Probabilistic Models
(Bayesian Networks)



# Turing Award 2011


Turing Award Citation:

"**Judea Pearl** is credited with the invention of Bayesian networks, a mathematical formalism for defining complex probability models, as well as the principal algorithms used for inference in these models. This work not only revolutionized the field of AI but also became an important tool for many other branches of engineering and the natural sciences."


Image Source: Lex Freidman Podcast

# Probability and Neural Nets

“a classy synthwave apartment on mars, digital art”

“an intricate line drawing of new your subway station full of trumpet players”

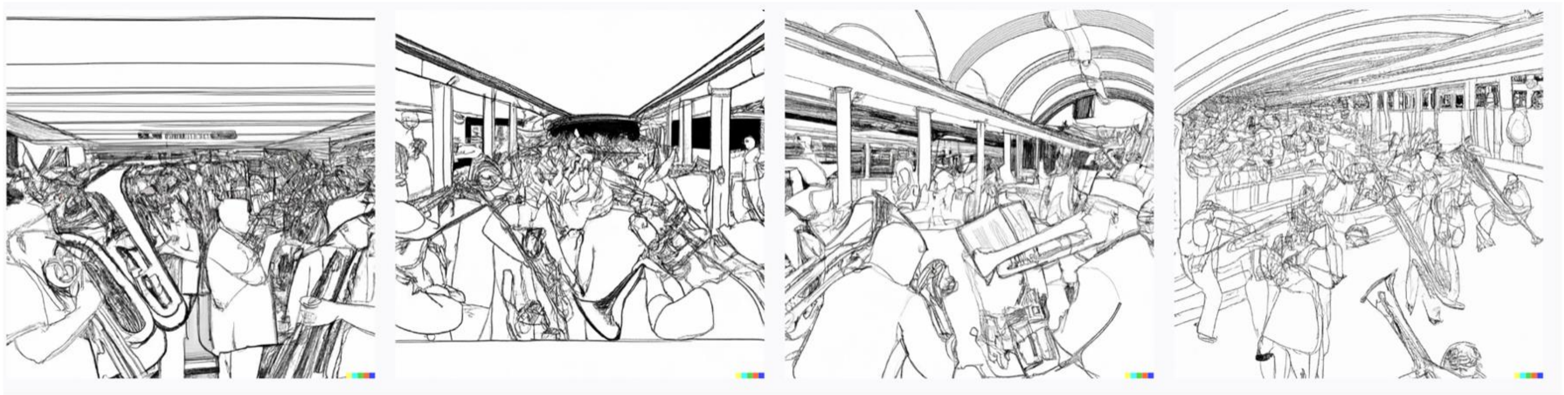
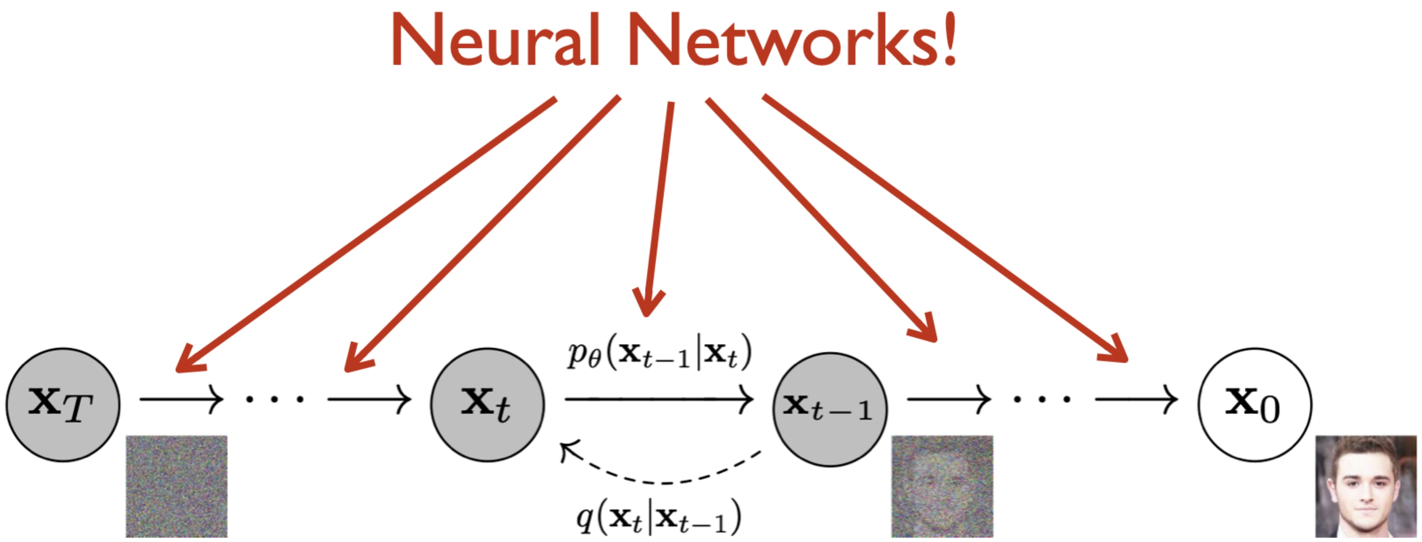


Image Source: Prof Berg-Kirkpatrick

# Probability and Neural Nets




$$\mathbb{E}_q\Big[\underbrace{D_{\mathrm{KL}}(q(\mathbf{x}_T|\mathbf{x}_0)\parallel p(\mathbf{x}_T))}_{L_T}+\sum_{t>1}\underbrace{D_{\mathrm{KL}}(q(\mathbf{x}_{t-1}|\mathbf{x}_t,\mathbf{x}_0)\parallel p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t))}_{L_{t-1}}\underbrace{-\log p_\theta(\mathbf{x}_0|\mathbf{x}_1)}_{L_0}\Big]$$

source: Ho et al. 2020

# Breakthrough in RL

## HOW THE ARTIFICIAL-INTELLIGENCE PROGRAM ALPHAZERO MASTERED ITS GAMES

**By James Somers**
December 28, 2018


*In 2016, a Google program soundly defeated Lee Sedol, the world's best Go player, in a match viewed by more than a hundred million people.* Photograph by Ahn Young-joon / AP

*Chess commentators have praised AlphaZero, declaring that the engine "plays like a human on fire."* Photograph Courtesy DeepMind Technologies

# Probability Review

# Probability in AI

Probability Theory == "How knowledge affects belief" (Poole and Mackworth)

What is the probability that it is raining out?

How much do I believe that it is raining out?

Viewing probability as measuring belief (rather than frequency of events) is known as the Bayesian view of probability (as opposed to the frequentist view).

# Discrete Random Variables

Discrete random variables, denoted with capital letters: e.g., $X$

Domain of possible values for a variable, denoted with lowercase letters: e.g., $\{x_1, x_2, x_3, \dots, x_n\}$

Example: Weather $W$ ; $\{w_1 = sunny, w_2 = cloudy\}$

# Unconditional (prior) Probability

$$P(X = x)$$

e.g., What is the probability that the weather is sunny?

$$P(W = w_1)$$

# Axioms of Probability

$$P(X = x) \geq 0$$

$$\sum_{i=1}^{n} P(X = x_i) = 1$$

$$P\left(X = x_i \; or \; X = x_j\right) = P(X = x_i) + P\left(X = x_j\right) \; if \; x_i \neq x_j$$

**Mutually Exclusive!**

# Conditional Probability

$$P(X = x_i | Y = y_j)$$

"What is my belief that $X = x_i$ if I already know $Y = y_j$"

Sometimes, knowing Y gives you information about X, i.e., changes your belief in X. In this case X and Y are said to be *dependent*.

$$P(X = x_i | Y = y_j) \neq P(X = x_i)$$

# Webclicker

Link: https://webclicker.web.app/

Login using UCSD Google account

If NO UCSD account –
Use a personal Google account

Course code: **KSALDG**

# Marginal Independence

Course code: KSALDG

$$P(X = x_i | Y = y_j) = P(X = x_i)$$

Sometimes knowing Y does not change your belief in X. In this case, X and Y are said to be independent.

$$P(W = w_i | Y = y_j) = P(W = w_i)$$

If $W$ denotes the weather today, For which variable Y is the above statement most likely true?

A. Y = The weather yesterday
B. Y = The day (Mon, Tue…) of the week
C. Y = The temperature

# More independence

Course code: KSALDG

Consider two students Roberto and Sabrina, who both took the same test.
Define the following random variables:

R = Roberto aced the test

S = Sabrina aced the test

*Assume both students have similar ability.*

What is the most logical relationship between $P(R = 1)$ and $P(R = 1|S = 1)$?

*A.* $P(R = 1) = P(R = 1|S = 1)$

*B.* $P(R = 1) > P(R = 1|S = 1)$

*C.* $P(R = 1) < P(R = 1|S = 1)$

# Conditional Independence

Course code: KSALDG

What if you also know the test was easy (variable T)?

A. $P(R = 1|T = 1) = P(R = 1|T = 1, S = 1)$

B. $P(R = 1|T = 1) > P(R = 1|T = 1, S = 1)$

C. $P(R = 1|T = 1) < P(R = 1|T = 1, S = 1)$

R and S are conditionally independent given T. I.e., if you already know T, knowing S does not give you additional information about R.

# More independence

Course code: KSALDG

Consider two events:

B = A burglar breaks into your apartment

E = An earthquake occurs

Are these events independent or dependent? (i.e., does knowing that one happened change your belief in the other?)

A. They are independent because knowing that one happened does not change your belief that the other happened.

B. They are dependent, because knowing that one happened changes your belief that the other happened.

# Conditional dependence

Course code: KSALDG

$$P(B=1) = P(B=1|E=1) = P(B=1|E=0)$$

Now consider a third event:

A = Your alarm goes off

Which of the following relationships best models beliefs about the world?

A. $P(B=1|A=1) = P(B=1|A=1, E=1)$

B. $P(B=1|A=1) > P(B=1|A=1, E=1)$

C. $P(B=1|A=1) < P(B=1|A=1, E=1)$

# That’s all folks!